

GenServer version 5.1.3  
Copyright (c) 1993 - 2003 Computer Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14 2003, 11:42:22 Search time: 4117 seconds  
(without alignments)  
11136.303 Million cell updates/sec

Title: US-09-910-428-1

Perfect score: 26

Sequence: 1 gtagctaatcttctctgtaccag 26

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues 32308152

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

- 1: em\_estla.\*
- 2: em\_estlum.\*
- 3: em\_estlin.\*
- 4: em\_estm.\*
- 5: em\_estec.\*
- 6: em\_estpl.\*
- 7: em\_estro.\*
- 8: em\_hic.\*
- 9: gb\_est1.\*
- 10: gb\_est2.\*
- 11: gb\_hic.\*
- 12: gb\_est3.\*
- 13: gb\_est4.\*
- 14: gb\_est5.\*
- 15: em\_estfun.\*
- 16: em\_estom.\*
- 17: gb\_gss.\*
- 18: em\_gss\_hum.\*
- 19: em\_gss\_inv.\*
- 20: em\_gss\_pla.\*
- 21: em\_gss\_vit.\*
- 22: em\_gss\_fun.\*
- 23: em\_gss\_mam.\*
- 24: em\_gss\_mus.\*
- 25: em\_gss\_other.\*
- 26: em\_gss\_pro.\*
- 27: em\_gss\_tod.\*

Prod % is the number of results predicted by Shanon's law, a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	19.8	76.2	458 14	BQ397777 NISC_mgc2
2	19.8	76.2	492 12	Bf513573 Cl H RMT
3	19.8	76.2	549 10	AW644499 cm41h03.w
4	19.8	76.2	566 12	Bf071855 db47d03.x
5	19.8	76.2	707 14	BQ188518 nt-F-ETI-
6	19.8	76.2	719 13	Bf089332 Bf089332

C	7	19.6	75.4	862 9	AA798980	AA798980 v95102.f
C	8	19.6	75.4	355 17	UN50281	AL231319 Tetradon
C	9	19.6	75.4	540 17	AV89840	AV89840 AV89840
C	10	19.2	73.8	803 13	B161824	B161824 603207531
C	11	19.2	73.8	1083 17	CNS038HP	AL236302 Tetradon
C	12	18.8	72.3	372 17	A2731725	A2731725 RPT 24-8
C	13	18.8	72.3	372 13	B1321744	B1321744 sal98d12
C	14	18.4	72.3	516 13	B1646299	B1646299 GMR30008B
C	15	18.4	72.3	626 13	B1646295	B1646295 GMR30007A
C	16	18.8	72.3	701 10	AM10404	AM10404 st35f01.x
C	17	18.8	72.3	796 17	HH718132	HH718132 HOMOG61TR
C	18	18.6	71.5	224 10	AM417202	AM417202 53310 MAR
C	19	18.6	71.5	271 9	AV131634	AV131634 AV131634
C	20	18.6	71.5	388 17	B63995	B63995 RPT-11-2111
C	21	18.6	71.5	471 13	B1898647	B1898647 479658 MA
C	22	18.6	71.5	476 13	BH483804	BH483804 537386 MA
C	23	18.6	71.5	508 10	AV663308	AV663308 AV663308
C	24	18.6	71.5	514 13	BM088926	BM088926 502504 MA
C	25	18.6	71.5	517 13	BM253554	BM253554 514670 MA
C	26	18.6	71.5	549 12	BE809805	BE809805 215575 MA
C	27	18.4	71.5	542 12	HA751000	HA751000 203696 MA
C	28	18.6	71.5	539 12	BP752343	BP752343 204711 MA
C	29	18.6	71.5	548 10	AM660850	AM660850 98343 MAR
C	30	18.6	71.5	558 10	AV291824	AV291824 AV591824
C	31	18.6	71.5	550 12	BE750323	BE750323 201653 MA
C	32	18.6	71.5	564 13	B1535110	B1535110 398443 MA
C	33	18.6	71.5	574 13	B1737379	B1737379 466471 MA
C	34	18.6	71.5	602 13	B1535654	B1535654 399324 MA
C	35	18.6	71.5	651 10	BH649754	BH649754 BH649754
C	36	18.6	71.5	768 17	AG097938	AG097938 Pan 11001
C	37	18.6	71.5	843 17	CNS07961	AL435400 17 end of
C	38	18.6	71.5	941 17	A2136491	A2136491 SP 0159 B
C	39	18.2	70.0	399 12	A2598392	A2598392 RPT 21-2
C	40	18.2	70.0	548 17	A2976761	A2976761 2M052822
C	41	18.2	70.0	564 17	A2928254	A2928254 RCT-21-2
C	42	18.2	70.0	567 12	BE940874	BE940874 EST420453
C	43	18.2	70.0	593 9	A1857121	A1857121 MENAD-23
C	44	18.2	70.0	632 17	BH761980	BH761980 BMR0356C
C	45	18.2	70.0	655 17	BH051439	BH051439 RPT-24-3

## ALIGNMENTS

RESULT 1  
BQ397777  
LOCUS  
DEFINITION  
NISC\_mgc2a02.g1 Scores NREG Xenopus laevis cDNA clone IMAGE:5161995  
BQ397777  
VERSION  
BQ397777.1 GI:21085464  
KEYWORDS  
SOURCE  
ORGANISM  
Xenopus laevis  
African clawed frog.  
REFERENCE  
NISC\_mgc2a02.g1 Scores NREG Xenopus laevis cDNA clone IMAGE:5161995  
TITLE  
NISC\_mgc2a02.g1 Scores NREG Xenopus laevis cDNA clone IMAGE:5161995  
JOURNAL  
COMMENT

458 bp mRNA linear EST 22-MAY-2002  
NISC\_mgc2a02.g1 Scores NREG Xenopus laevis cDNA clone IMAGE:5161995  
BQ397777  
VERSION  
BQ397777.1 GI:21085464  
KEYWORDS  
SOURCE  
ORGANISM  
Xenopus laevis  
African clawed frog.  
REFERENCE  
NISC\_mgc2a02.g1 Scores NREG Xenopus laevis cDNA clone IMAGE:5161995  
TITLE  
NISC\_mgc2a02.g1 Scores NREG Xenopus laevis cDNA clone IMAGE:5161995  
JOURNAL  
COMMENT



BACKWARD: CAGGAAACGCTATGACC  
 Plate: 0140 Row: H Column: 03  
 Seq primer: T7 primer.  
 Location/Qualifiers  
 1..549  
 /organism="Xenopus laevis"  
 /db\_xref="taxon:8355"  
 /clone="PBX0140H03"  
 /cdate="11/11/96"  
 /sex="female"  
 /library="unfertilized egg"  
 /dev\_stage="unfertilized egg"  
 /lab\_host="DH10B"

BASE COUNT 179 a 86 c 112 g 172 t  
 ORIGIN  
 Query Match 76.2% Score 19.8; DB 10; Length 549;  
 Best Local Similarity 31.3% Fred. No. 4.3e+02;  
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 1 CTCCTCAATCTTTTCGTAC 23  
 11111111111111111111  
 242 GTGCTCAATCTTTTCGTAC 264

RESULT 4  
 BP071855/c  
 LOCUS 566 bp mRNA linear EST 17-OCT-2000  
 DEFINITION B447403 x1 B1a-kshear/Soares normalized Xenopus egg library X-107-05  
 laevis cDNA clone IMAGE:3301637 3', mRNA sequence.  
 ACCESSION BP071855  
 KEYWORDS BP071855.1 GI:10848494  
 EST  
 SOURCE African clawed frog.  
 ORGANISM Xenopus laevis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Anura; Mesobatrachia; Pipridae; Pipidae;  
 Xenopodidae; Xenopus.  
 1 (bases 1 to 566)  
 1 CTTTTCCTCAATCTTTTCGTAC 23  
 11111111111111111111  
 242 GTGCTCAATCTTTTCGTAC 264

REFERENCE  
 1. Clifton S., Johnson S. I., Rumberg R., Song T., Hillier, L., Page, P., Martin, J., Wylie, T., Underwood, K., Theisling, B., Bowers, Y., Peterson, B., Gibbons, M., Harvey, N., Riller, E., Jackson, Y., McCann, R., Waterston, R., and Wilson, R.  
 Washu Xenopus EST Project, 1999  
 Unpublished (1999)  
 Contact: Sandy Clifton, Ph.D.  
 Washu Xenopus EST Project, 1999  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@wustl.edu  
 Library constructed by Bento Soares and M. Fatima Bonaldo  
 (University of Iowa). DNA Sequencing by: Washington University  
 Genome Sequencing Center  
 Clone Distribution: Xenopus clones from this library are available  
 through the I.M.A.G.E. Consortium/ILM at: image.llnl.gov  
 Seq primer: -400P from Gibco  
 High quality sequence stop: 458.

FEATURES  
 source

Location/Qualifiers  
 1..566  
 /organism="Xenopus laevis"  
 /db\_xref="taxon:8355"  
 /clone="IMAGE:3301637"  
 /cdate="11/11/96"  
 /sex="female"  
 /library="unfertilized egg"  
 /dev\_stage="unfertilized egg"  
 /lab\_host="DH10B"

BASE COUNT 180 a 114 c 85 g 187 t  
 ORIGIN  
 Query Match 76.2% Score 19.8; DB 12; Length 566;  
 Best Local Similarity 31.3% Fred. No. 4.3e+02;  
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 1 CTCCTCAATCTTTTCGTAC 23  
 11111111111111111111  
 242 GTGCTCAATCTTTTCGTAC 292

RESULT 5  
 B0188518/c  
 LOCUS 707 bp mRNA linear EST 30-APR-2002  
 DEFINITION U1-E-EJ1-a1v-h-18-0-U1.r1 U1-E-EJ1 Homo sapiens cDNA clone  
 U1-E-EJ1-a1v-h-18-0-U1 5', mRNA sequence.  
 ACCESSION B0188518  
 KEYWORDS B0188518 GI:20364059  
 EST  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 707)  
 1 B0188518 M.F., Lennon, G. and Soares, M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 Genome Res. 6 (9), 791-806 (1996)  
 97044477  
 Contact: Soares, MB  
 Program for Rat Gene Discovery and Mapping  
 University of Iowa  
 451 Rockness Medical Research Building Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: msoares@blue.weeg.uiowa.edu  
 Tissue Procurement: Dr. Gregg Hageman  
 cDNA library preparation: Dr. M. Bento Soares, University of Iowa  
 DNA library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.research.com).  
 Seq primer: M13 REVERSE.

FEATURES  
 source

Location/Qualifiers  
 1..707  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="U1-E-EJ1-a1v-h-18-0-U1"



Db 626 GCGCCGACGCTCTCTGACACG 601

RESULT 8  
CNS02ZRI 905 bp DNA linear GSS 15-MAY-2000  
LOCUS Tetradodon nigroviridis genome survey sequence TUC-01 end of clone  
DEFINITION 183001 of library G from Tetradodon nigroviridis, genomic survey  
sequence.

ACCESSION  
VERSION AL221319.1 GI:7880138  
KEYWORDS GSS; genome survey sequence.  
SOURCE Tetradodon nigroviridis.  
ORGANISM Tetradodon nigroviridis  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorphi; Tetradodontiformes;  
Tetradodontidae; Tetradodon.

REFERENCE  
AUTHORS Roest-Crollius, H., Jallou, O., Dasilva, C., Bonneau, L., Fisher, C.,  
Bernol, A., Fizames, C., Wincker, P., Brottier, F., Quetier, F.,  
Saurin, W. and Weissenbach, J.  
TITLE Human gene number estimate provided by genome wide analysis using  
Tetradodon nigroviridis DNA sequence  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 905);  
AUTHORS Roest-Crollius, H., Jallou, O., Dasilva, C., Fizames, C., Fisher, C.,  
Bonneau, L., Billault, A., Quetier, F., Saurin, W., Bernol, A. and  
Weissenbach, J.  
TITLE Characterization and repeat analysis of the compact genome of the  
freshwater pufferfish Tetradodon nigroviridis  
JOURNAL Unpublished  
REFERENCE 3 (bases 1 to 905)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (12-APR-2000)  
COMMENT This sequence is a single read and was generated as part of a large  
scale clone and sequencing project of the Tetradodon nigroviridis  
genome. For more information, please take a look at  
<http://www.genoscope.cns.fr/Tetradodon>

FEATURES  
SOURCE  
LOCATION/Qualifiers  
1..905  
/organism="Tetradodon nigroviridis"  
/db\_xref="taxon:99883"  
/clone="183001"  
/clone\_1lb="G"  
/note="Genoscope sequence ID: CNS02ZRI01SPI-end"  
pur="ori"

BASE COUNT 212 a 203 c 226 g 262 t 2 others

ORIGIN

Query Match 75.4%, Score 19.6; DB 17; Length 905;  
Best Local Similarity 84.5%; Pred. No. 6, le-02;  
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 252 CTGCTTAACTCTCTGACACG 277

RESULT 9  
AVB89040 580 bp mRNA linear EST 09-NOV-2001  
LOCUS AVB89040 Nori Satoh unpublished cDNA library, cleavage stage embryo  
DEFINITION Clona intestinalis cDNA clone rcic31104 3', mRNA sequence.  
ACCESSION AVB89040  
VERSION AVB89040.1 GI:16877755  
KEYWORDS EST.  
SOURCE Clona intestinalis.  
ORGANISM Clona intestinalis  
Eukaryota; Metazoa; Chordata; Trochordata; Ascidiacea; Enteropneusta;  
Phlebobranchia; Clonidae; Clona

AUTHORS Satoh, N., Satou, Y., Kohara, Y. and Shin, I.  
TITLE Expressed genes in Clona intestinalis  
JOURNAL Unpublished (2000)  
COMMENT Contact: Nori Satoh  
Department of Zoology  
Kyoto University  
Sakyo-ku, Kyoto 606-8502, Japan  
Tel: 81-75-753-4081  
Fax: 81-75-705-1113  
E-mail: sat@ksc.kyoto-u.ac.jp

FEATURES  
SOURCE  
LOCATION/Qualifiers  
1..580  
/organism="Clona intestinalis"  
/db\_xref="taxon:7719"  
/clone="rcic31104"  
/clone\_1lb="Nori Satoh unpublished cDNA library, cleavage  
stage embryo"  
/tissue\_type="whole animal"  
/dev\_stage="cleavage stage embryo"  
/note="Vector: pBluescript SK"

BASE COUNT 199 a 98 c 94 g 188 t 1 others

ORIGIN

Query Match 73.8%; Score 19.2; DB 10; Length 580;  
Best Local Similarity 87.5%; Pred. No. 7, 9e-02;  
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 290 GCGCGAAGGATTCGCGACCA 313

RESULT 10  
B1461824 803 bp mRNA linear EST 21-AUG-2001  
LOCUS B1461824  
DEFINITION M3207531F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:5274287 5',  
mRNA sequence.  
ACCESSION B1461824  
VERSION B1461824.1 GI:15252480  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 803)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@strausberg.nih.gov  
Tissue Procurement: Miklos Palokovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NIH/RI), Shitaki  
Toshiyuki and Piero Garlundi (RIKEN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/INT at:  
<http://image.llnl.gov>  
Plate: LLAM1689 row: n column: 08  
High quality sequence stop: 761.  
LOCATION/Qualifiers  
1..803  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5273287"  
/clone\_1lb="NIH\_MGC\_97"  
/lab\_host="DH10B"  
/note="Organ. testis, Vector: pBluescript (modified  
pBluescript KS+); Site\_1: BamHI; Site\_2: SalI; XhoI (quecga  
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTNN 3',  
size-selected 1cf average insert size 2.2 kb and  
normalized to KOF 5. This is a primary library enriched  
for full-length clones and constructed using the





Host Local Similarity: 90.98% Prod. No.: 1.1e+03;  
Matches: 291 Conserved: 0; Mismatches: 2; Indels: 0; Gaps: 0.

Y 5 TCTAATCTTTTCGTACACAG 26  
|||||  
db 170 TCTAATCTTTTCGTACACAG 191

## RESULT 15

LENS

61969065 626 bp mRNA Linear EST 21.07.2001  
DEFINITION GmR1083 Glycine max cDNA clone Gm-R1083-2360 J.

Accession: GmR1083

Version: 1.0

Keywords:

Source:

Organism:

glycine max

Accession: GmR1083

Version: 1.0

Keywords:

Source:

Organism:

glycine max

Accession: GmR1083

Version: 1.0

Keywords:

Source:

Organism:

glycine max

Accession: GmR1083

Version: 1.0

Keywords:

Source:

Organism:

glycine max

Accession: GmR1083

Version: 1.0

Keywords:

Source:

Organism:

glycine max

Accession: GmR1083

Version: 1.0

Keywords:

Source:

Organism:

glycine max

Accession: GmR1083

Version: 1.0

Keywords:

Source:

Organism:

glycine max

Accession: GmR1083

Version: 1.0

Keywords:

Source:

Organism:

glycine max

Accession: GmR1083

Version: 1.0

GmR1083 library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the Incyte Genomics clone ID of the original cDNA library that is also listed under 'OTHER EST'.

RASH COUNT 193 a 158 c 96 g 156 t 23 others

Query Match 72.4% Score 18.8; DB 13; Length 626;

Host Local Similarity: 90.98% Prod. No.: 1.2e+03;

Matches: 291 Conserved: 0; Mismatches: 2; Indels: 0; Gaps: 0;

Y 5 TCTAATCTTTTCGTACACAG 26

|||||

db 251 TCTAATCTTTTCGTACACAG 272

Search completed: January 14, 2003, 15:10:27  
Job time: 41.8117 secs

## FEATURES

SOURCE

1..626  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="Gm-R1083-2360"  
/clone\_lib="Gm-R1083"  
/note="The library Gm-R1083 is a sequence-driven, retracted set of 4,992 clones selected from cDNA libraries from various tissues and stages of development of soybean. It represents 117 sequences from the progenitor library Gm-R1083 (from mature roots of 2 month old greenhouse grown 'Williams' soybean plants); 820 sequences from the progenitor library Gm-R1083 (from 2 to 3 week old whole plants of Williams); and 1055 sequences from library Gm-R1083 (from 'Superior' plants whose seedlings were inoculated with Bradyrhizobium japonicum, courtesy of Dr. Gary Stacey). The 5' ESTs of the source clones from the different progenitor libraries was used to select singletons, or a representative of each contig, which were retracted to form library Gm-R1083. The cDNA clones of the retracted Gm-R1083 library were then sequenced at the 3' end. The contig analysis to select unique genes was performed by the Laboratory of Ernest Rietzel, Center for Computational Genomics and Bioinformatics, University of Minnesota. <http://web.its.umn.edu/bioinformatics/>. Retacking was performed by Incyte Genomics, St. Louis, <http://www.incyte.com>, and 5' sequences by the Keck Center for Comparative and Functional Genomics, University of Illinois, <http://www.its.uiuc.edu/keck/5.1.2.html>.  
Note: The corresponding 5' EST from each clone in the

Seq primer: 5' TTTTCTTTTCTTTTCTTTT(A/C/G)-3'

Location/Qualifiers